SUMMARY

This report describes research on the genetic status and relationships among coastal California salmonid populations. The scope of work broadened from the original contract investigating population structure and genetic diversity of coho populations to include research on steelhead and chinook populations and the development of a Geographical Information System (GIS).

Substantial progress was made in documenting coho population genetic diversity within the California Central Coastal (CCC) ESU. A suite of highly polymorphic microsatellite DNA markers was identified and used to establish genetic diversity within and among 57 collections of coho salmon from 14 watersheds. The samples encompass the southern end of the Southern Oregon / Northern California (SO/NC) ESU, the entire CCC ESU, and the South of San Francisco (SSF) ESU recognized by California's Endangered Species Act. Genetic distances among samples support the present State of California ESU structure, forming statistically significant clusters of samples corresponding to the CCC, the SSF, and the most southerly of samples from the SO/NC ESU (Eel and Mattole Rivers). Samples from the Klamath and Trinity Rivers are significantly separated from the Eel / Mattole River samples and from the CCC and SSF ESUs clusters. Sampling of different year-classes at seven sites reveals that temporal variation is typically significant, though smaller than the geographic component of population genetic structure. The congruence of genetic and geographic distance is surprising in light of the history of coho stock transfers within California and between California and other Pacific Coast states. Stock transfers appear to have left no genetic mark on extant populations. Alternatively, or in addition to stock transfers, the diversifying effects of genetic drift within the relict coho populations of California may be keeping pace with whatever homogenization has been or is being effected by hatchery practices.

We find many significant deviations between observed genotypic composition of coho salmon populations and the composition expected under random mating. These deviations occur both in juvenile samples, in which they might be expected, owing to kinship among individuals, and in adult samples, in which they are not expected based on the population genetic literature for natural populations of Pacific salmon. We discriminate and attempt to correct for the contributions of two different potential causes of deviations from random mating equilibrium – admixture in collections of individuals from genetically differentiated subpopulations and kinship. Partitioning samples based on independent biological information (sex, size, date caught, precise site of collection, whether marked, type of mark) does successfully reduce the deviations within some samples.

In most juvenile samples, many pairs of individuals show statistically significant odds of being full brothers and sisters. Because such samples yield biased and inaccurate estimates of the genetic diversity in the adult spawning population, population geneticists in the past have avoided using juvenile samples. Nevertheless, the depressed state of coho salmon populations often precludes collections of sufficient numbers of adults. Juveniles, on the other hand, are more readily available in large numbers. Of the 57 collections available for this study, 27 comprised juveniles. To salvage these important samples for genetic analysis, we apply methods pioneered in our lab for adjusting samples for family structure to derive unbiased and accurate estimates of adult allele frequencies. Related individuals are either removed and replaced with reconstructed parents or simply removed from a sample, resulting in a sample that is smaller but

usually closer to, if not in random mating equilibrium. Nearly half of the samples used to infer the geographic distribution of genetic diversity in this study are adjusted juvenile samples.

A large fraction of coho samples continues to deviate from random mating expectations after adjusting samples for substructure and kinship. Deviations from random mating proportions in some adult samples could be explained by inbreeding, and a significant excess of individuals homozygous for multiple markers supports this hypothesis. The non-equilibrium state of coho juveniles from Green Valley Creek and their highly aberrant genetic distance to other populations in the CCC ESU is of special concern, as fish from this population are currently being reared for a hatchery-based recovery effort in the Russian River watershed.

In order to estimate the genetic affinities of chinook salmon in the Russian River with other stocks in California, we examined seven DNA markers in 449 fish from the Russian and Eel Rivers and Lagunitas Creek. Genetic distances show that chinook salmon in the Russian River are distinct from those in the Eel and Klamath Rivers to which they are more closely related than chinook from the Central Valley of California.

Jason Watters, a Ph.D. student supported by this contract, examined the development and maintenance of alternative male phenotypes in coho salmon. He showed that the phenotypes of juvenile coho males are affected by rearing habitat and alternative male phenotypes have different reproductive success. Thus, the maintenance of alternative male phenotypes in wild spawning populations could be critical to population viability.

Finally, a web-based GIS that focuses on coastal near-shore processes and allows linkages and integration of marine and coastal stream environmental data was developed. It is the first GIS model to incorporate real-time ocean surface currents measurements derived from coastal high frequency radar stations. This web-based GIS has the potential to deliver up-to-date information to a broad audience in a timely manner. Custom PERL programming scripts were developed in collaboration with the REGIS laboratory at UC Berkeley. A CD Rom containing the database files, software, directory structure and scripts is included with this report.

INTRODUCTION

This report describes research done under contract #TW 99/00-110 a continuation of work initiated under contract #TW 96/97-10 from the Sonoma County Water Agency. The first contract was focused on the population genetics of coho salmon (*O. kisutch*) in Northern California, and this continued to be the major emphasis under the second contract. The scope of research on the second contract was expanded, however, to include research on life history variation in coho salmon as well as on the population genetics of steelhead (*O. mykiss*) and chinook salmon (*O. tshawytscha*). We proposed, moreover, to develop a geographical information system (GIS), to enable synthesis and visualization of environmental and genetic data critical to management of coastal salmonid resources. Progress towards achieving the specific tasks is summarized in the body of this report.

Salmonid conservation requires identification of appropriate management units in a complex, geographically structured hierarchy of populations. Population genetics documents biodiversity at various levels in a population hierarchy and provides a variety of tools for resource management. In the first contract, for example, we developed molecular diagnostic tests that discriminates steelhead, coho and chinook salmon, which co-occur in juvenile and carcass samples and can be difficult to distinguish morphologically (Greig et al. 2002). Within Pacific salmon species, the challenge is to identify how geographically structured biodiversity is influenced by hatcheries, environmental degradation, and ocean harvests. Finally, at the level of the local spawning run, estimates of effective population size (N_e) from genetic data can help predict the rate of loss of biodiversity and identify foci for recovery efforts. All of these genetic measures are essential components of viable population size (VP) estimates, which are central to management and restoration efforts.

Genetics of geographically structured populations

A brief review of basic population genetic principles will aid in understanding of some exceptional findings to be presented in this report. One of the oldest principles of population genetics, named, after its co-discoverers, the Hardy-Weinberg Principle (Hedrick 2000), describes the expected proportion of genotypes in a randomly mating population. If a hypothetical gene (or **locus**) has two alleles in a population, A_1 and A_2, with relative frequencies of p and $q=(1-p)$, respectively, then the proportions among N adults of the three possible genotypes at this locus are given by the binomial expansion, $(p+q)^2 = Np^2 + 2Npq + Nq^2$. For example, if alleles A_1 and A_2 have frequencies of 0.7 and 0.3, respectively, then among 100 individuals in a sample of adults, we expect to find 49 A_1A_1 homozygotes, 42 A_1A_2 heterozygotes, and 9 A_2A_2 homozygotes. Populations conforming to this principle are said to be in Hardy-Weinberg (H-W) or random mating equilibrium. The H-W Principle, which is easily extended to the multiple alleles typical of the highly polymorphic microsatellite DNA markers used in this research, simplifies enormously the description of populations, reducing the number of parameters to n alleles per locus, rather than the $n(n + 1)$ genotypes formed by sexual reproduction of diploid organisms.

The significance of differences between the observed and expected proportions of genotypes in populations can be tested in a number of ways, classically by a goodness-of-fit χ^2-test but, more recently, by Fisher exact tests, Markov-Chain approximations of the exact test, or permutation tests, which are more appropriate to the small expected numbers generated by many low

frequency alleles. The vast literature on the genetics of Pacific salmon populations shows that natural populations generally conform to the Hardy-Weinberg Principle (*e.g.* Bartley et al 1992a, b), implying that mating is more or less at random among spawning adults. Here, we report many exceptions to random mating equilibrium.

The principle of random mating equilibrium can be extended to multiple genes considered simultaneously. For example, with two genes, A and B, each with two alleles, A_1, A_2 and B_1, B_2, the expected proportion of each gamete at random mating equilibrium can be calculated as the product of the relevant allelic frequencies, *e.g.* the expected frequency of an egg carrying the A_1B_2 combination is the product pr, if the relative frequencies of A_1 and B_2 are p and r, respectively. As for the single-locus equilibrium described by the Hardy-Weinberg Principle, statistical tests of departures in samples from random multi-loci associations of alleles into gametes can be made, usually for pairwise combinations of markers. These tests are commonly called tests of **linkage disequilibrium** or **LD** (though, since physical linkage is not required, they are more properly called tests of gametic-phase disequilibrium; Hedrick 2000). Again, Pacific salmon populations are generally in gametic-phase equilibrium, but we report many exceptions here.

A number of factors can cause deviations from random mating expectations. In order to understand these and the results to be presented in this report, we must first consider how the genetic diversity of a species can be partitioned into components within and among population units, ranging from local, randomly mating populations (or **demes**) to subpopulations to the total species. Wright (1931, 1943) partitioned genetic variation within a species, using F-statistics, which measure the average genetic correlation between pairs of gametes derived from different levels in a population hierarchy. At the basal level of this hierarchy, the correlation between gametes drawn from different individuals within a deme is symbolized as F_{IS}. F_{IS} is zero in a randomly mating subpopulation but is positive when there are excesses of homozygotes relative to H-W expectations. Inbreeding, mating among related individuals, causes excesses of homozygotes and deficiencies of heterozygotes, in which case F_{IS} is positive.

If a species is subdivided into partially isolated, finite subpopulations, mating among individuals in the total population cannot take place at random and there will be genetic drift within each subpopulation. The effect on the proportion of genotypes in the species is analogous to the effect of inbreeding: local populations will tend towards fixation, with a decline in heterozygosity, but genetic diversity will be preserved among rather than within subpopulations. The genetic correlation between gametes drawn from different demes or subpopulations, with respect to allelic frequencies in the total population, is given by F_{ST}, the ratio of the variance of allelic frequencies among subpopulations to the maximum. When local populations diverge from one another, there will be an excess of homozygotes and a deficiency of heterozygotes, with respect to random mating expectations, summing across subpopulations. The principle is readily understood at the extreme, in which each subpopulation is fixed for one allele or another ($F_{ST}=1.0$); in this case, there are no heterozygotes in the total population. Heterozygote deficiency can result artificially from the unwitting admixture, in collections from natural populations, of individuals from genetically differentiated demes. This artificially induced deficiency of heterozygotes, which is known as the **Wahlund effect**, after its discoverer (Hedrick 2000), will be illustrated in the study of coho salmon reported here.

Genetics of juvenile salmon populations

Finally, we consider the consequences of sampling juveniles rather than adults for studies of genetic diversity. The old and very sound advice for students of salmon population genetics is to avoid sampling juveniles:

> "The correct way of approaching the question of possible genetic differences between subpopulations is to sample the spawners. ...it is dangerous to draw conclusions about reproductive isolation between adults by estimating allelic frequencies in their progeny. Differences caused by a small number of reproducing adults without any reproductive isolation can become highly statistically significant when a large number of progeny are sampled." (Allendorf and Phelps 1981).

Nevertheless, the presently depressed state of salmon populations, particularly coho salmon populations in Central and Northern California, often precludes the collection of a sufficient number of adult samples for genetic analysis. Juveniles, either fry or smolts, which are more easily collected in large numbers, are often the only sample available. In such samples, however, the H-W Principle does not apply. Either excesses or deficiencies of heterozygotes with respect to random mating expectations can occur, depending on the number and sizes of families present in a juvenile collection and on the genotypes of their parents. Likewise, linkage disequilibrium can often be generated, owing to the limited number of gametic combinations passed to progeny from a small number of parents; indeed, linkage disequilibrium provides a sensitive indicator of family structure in juvenile samples. Departures from random mating equilibrium will be illustrated for juvenile samples of coho salmon. We have endeavored to correct the allele-frequency estimates for the family structure in these samples, following the approach pioneered by us previously (Banks et al. 2000), thus to salvage these samples for use in our study of coho salmon diversity.

POPULATION GENETICS OF COASTAL CALIFORNIA COHO SALMON POPULATIONS

Introduction

The specific tasks in our scope for work were: 1) to determine relatedness in samples comprised of juveniles, 2) to determine temporal genetic variation among year classes, 3) to estimate genetic divergence among and effective population sizes of spawning runs, 4) to determine genetic change between historical and extant coho populations, to assess influence of hatchery plantings and reductions in abundance, 5) to relate the genetic diversity of California coho populations to environmental and biological factors being measured in the sampling process.

The contract also supported Kate Bucklin's doctoral thesis research on nucleotide sequence diversity and phylogeny across the North Pacific range of coho salmon. However, as described in the annual report for 2001, so little variation was detected at the nucleotide level that this research was not pursued and no results are presented here.

The major objective of this contract and its predecessor was to describe the genetic diversity of coho salmon populations along the central and northern coast of California, using highly polymorphic microsatellite DNA markers. Genetic diversity of coho salmon in this region was previously examined using protein markers, which have low levels of polymorphism and reveal little geographic structure (Bartley et al. 1992a). For our analysis, we selected seven

microsatellite DNA markers for their variability and apparent diversity among populations. The geographic coverage of our samples extends from the Klamath River, Del Norte Co., to Scott Creek, Santa Cruz Co., and includes populations from three Evolutionary Significant Units (ESUs), the Southern Oregon / Northern California ESU, the Central California ESU, and the South of San Francisco ESU, which the State California distinguishes from their Central California ESU. We present results for seven DNA markers, in over 1600 fish from 57 populations of coho salmon. These genetic data provide a context for understanding Sonoma County coho populations.

Materials and Methods

Microsatellite DNA markers
An extensive survey of known salmonid microsatellite DNA markers established a suite for assessing genetic diversity of California coho salmon. Investigation into published primers for the six Pacific species produced 67 microsatellites for testing. The screening processes used samples from Scott Creek (Santa Cruz County), Noyo River (Mendocino County), Eel River (Humboldt County) and Smith River (Del Norte County) to examine variability and assess potential diagnostic power. Sixty-five microsatellites were eliminated leaving seven polymorphic, potentially diagnostic loci (Table 1). Multiplexing the seven microsatellites into three PCR reactions increased efficiency. The microsatellite *iso-Ots-2* is known to have species-specific differences and was included to ensure species identity (Greig et al 2002).

Table 1. Summary of microsatellites examined from six Pacific salmon and other species. Microsatellite screening results are coded as follows: (N) total number examined, (In Use) selected for use assessing populations in California, (ND) not diagnostic in California, (NV) not variable, fewer than 4 alleles, (NW) primers did not work.

A. Markers screened.

Species	N	In Use	ND	NV	NW
Oncorhynchus gorbuscha	7	0	0	5	2
O. keta	3	0	0	1	2
O. kisutch	13	1	2	4	6
O. mykiss	3	0	0	1	2
O. nerka	18	1	0	9	8
O. tshawytscha	22	4	3	10	5
Other	3	1	0	1	1
Total	69	7	5	31	26

B. Markers selected for use.

Microsatellite	Repeat #	# Alleles	Reference
Ots-2	Di	8	Banks et al. 1999
iso-Ots-2	Di	16	Greig et al. 2001
Ots-3	Di	12	Banks et al. 1999
Ots-103	Tetra	35	Nelson and Beacham 1999
Oki-1	Tetra	13	Smith, C. T et al. (1998).
One-13	Di	17	Scribner et al. 1996
P-53	Di	10	Park et al. 1996

Figure 1. Map of Northern California, showing watersheds and in-stream sites, from which coho salmon were collected for population genetic analysis.

Table 2. Samples of coho salmon used for genetic analysis. Stages are A= adults, S= smolts, Y= young of the year. Populations are designated by their Name codes in subsequent tables and figures. The criteria for subdividing collections from certain sites or drainages are listed.

Watershed	Tributary or Site	No.	Stage	Yr. Coll.	Name code	Criteria; Collectors
Klamath River	Iron Gate Hatchery	11	A	97/98	KIGHA97a	Ad clip, FL>56cm; CDFG
Klamath River	Iron Gate Hatchery	15	A	97/98	KIGHA97j	FL<56cm; CDFG
Klamath River	Iron Gate Hatchery	36	A	97/98	KIGHA97ll	Left clip, FL>56cm; CDFG
Klamath River	Iron Gate Hatchery	19	A	97/98	KIGHA97nl	No clip, FL>56cm; CDFG
Trinity River	Trinity River Hatchery	17	A	97/98	TRHA97s	FL<45cm; CDFG
Trinity River	Trinity River Hatchery	77	A	97/98	TRHA97l	FL>53cm; CDFG
Little River (Humboldt Co,)	Little River Delta	85	S	2000	LRS00-1	4/3/00-5/6/00; Simpson Timber Co.
Little River (Humboldt Co.)	Little River Delta	11	S	2000	LRS00-2	5/19/00-5/29/00; Simpson Timber Co.
SF Eel River	Hollowtree Creek	16	A	97/98	EHOLA97	Salmon Trawlers Assoc.
SF Eel River	Redwood Creek	92	S	97	EREDS97	Eel River Salmon Restoration Project
SF Eel River	Redwood Creek	22	A	98/99	EREDA98	Eel River Salmon Restoration Project
SF Eel River	South Fork Sproul Creek	34	S	1999	ESPRS99	Eel River Salmon Restoration Project
Mattole River	Mattole River Delta	75	S	98	MATS98-1	5/7/98 and 5/12/98; Mattole Salmon Group
Mattole River	Mattole River Delta	21	S	98	MATS98-2	5/19/1998; Mattole Salmon Group
Pudding Creek	Pudding Creek	33	Y	98	PUDY98h	9/23/1998; Georgia Pacific
Pudding Creek	Pudding Creek	43	Y	98	PUDY98k	10/27/1998; CDFG
Pudding Creek	Upper Pudding Creek	4	Y	98	PUDY98u	9/23/1998; Georgia Pacific
SF Noyo River	Egg Taking Station	47	A	97/98	NOYA97	Bill Cox, CDFG
SF Noyo River	Egg Taking Station	47	A	99/00	NOYA99	CDFG
Albion River	Albion Mainstem	23	A	98/99	ALBA98	Mendocino Redwood Co.
Albion River	Marsh Creek	18	Y	98	ALBY98	CDFG
Russian River	Warm Springs Hatchery	33	A	95/96	RRHA95	CDFG
Russian River	Warm Springs Hatchery	25	A	96/97	RRHA96	CDFG
Russian River	Warm Springs Hatchery	7	Y	97	RRHY97	CDFG
Russian River	Green Valley 10/27	9	Y	97	RRGV97	Michael Fawcett
Russian River	Green Valley	70	Y	98	RRGV98a	Michael Fawcett / SCWA
Russian River	Green Valley10/13	58	Y	98	RRGV98b	Michael Fawcett / SCWA
Russian River	Green Valley	8	Y	99/00	RRGV00	Michael Fawcett / SCWA
Russian River	Delta	8	S	97	RRDS97	Michael Fawcett / SCWA

(Table continues next page)

Table 2. Samples of coho salmon used for genetic analysis. Stages are A= adults, S= smolts, Y= young of the year. Populations are designated by their Name codes in subsequent tables and figures. The criteria for subdividing collections from certain sites or drainages are listed.

Watershed	Tributary or Site	No.	Stage	Yr. Coll.	Name code	Criteria; Collectors
Russian River	Delta	3	S	98	RRDS98	Michael Fawcett / SCWA
Russian River	Mirabel	2	Y	98	RRM98	Michael Fawcett / SCWA
Lagunitas Creek	Lagunitas Mainstem	9	A	96/97	LAGA96	Trihey Associates
Lagunitas Creek	Lagunitas Mainstem	7	A	97/98	LAGA97	CDFG
Lagunitas Creek	Lagunitas Mainstem	2	A	99/00	LAGA99	MMWD
Lagunitas Creek	Devils Gulch	9	A	96/97	LDGA96	Volunteers
Lagunitas Creek	Devils Gulch	10	A	97/98	LDGA97	Bob Chamberlain, Volunteers
Lagunitas Creek	San Geronimo	32	A	95/96	LSGA95	Bob Chamberlain, Volunteers
Lagunitas Creek	San Geronimo	19	A	96/97	LSGA96	Bob Chamberlain, Volunteers
Lagunitas Creek	San Geronimo	52	A	96/97	LSGA96	Bob Chamberlain, Volunteers
Lagunitas Creek	San Geronimo	61	A	97/98	LSGA97	Bob Chamberlain, Volunteers
Lagunitas Creek	San Geronimo	10	Y	96	LSGY96	Bob Chamberlain, Volunteers
Lagunitas Creek	San Geronimo	12	Y	98	LSGY98	Bob Chamberlain, Volunteers
Lagunitas Creek	San Geronimo Arroyo	36	A	96/97	LSGAA96	Bob Chamberlain, Volunteers
Lagunitas Creek	San Geronimo Arroyo	3	A	97/98	LSGAA97	Bob Chamberlain, Volunteers
Lagunitas Creek	San Geronimo Arroyo	21	Y	98	LSGAY98	Bob Chamberlain, Spring Class
Olema Creek	Mainstem	71(2X)	A	96/97	OLEA96	Natl. Park Service
Olema Creek	Mainstem	34 (2X)	A	97/98	OLEA97	Tomales Bay Assoc.
Olema Creek	Mainstem, Blueline	88	Y	98	OLEY98	Natl. Park Service
Redwood Ck. (Marin)	Mainstem	15	A	97/98	RWMA97	Natl. Park Service
Redwood Ck. (Marin)	Mainstem	24	Y	98	RWMY98	Jerry Smith
Waddell Creek	Mainstem	42	Y	99	WADY99low	RM 3.1 and 3.9; Jerry Smith
Waddell Creek	Mainstem	17	Y	99	WADY99up	RM 4.7; Jerry Smith
Scott Creek	Hatchery	43	A	95/96	SCA95	Monterey Bay Trout and Salmon Project
Scott Creek	Hatchery	57	A	97/98	SCA97	Monterey Bay Trout and Salmon Project
Scott Creek	Hatchery	42	A	98/99	SCA98	Monterey Bay Trout and Salmon Project
Scott Creek	Mainstem, Big & Mill Creeks	40	Y	99	SCY99low	RM 2.55, 3.55, B&M Cks.; Jerry Smith
Scott Creek	Mainstem, Upper Fork	20	Y	99	SCY99up	RM 4.9, Upper Fork; Jerry Smith